THE MODULATION OF FORMANT AMPLITUDE DURING ASSIMILATION PROCESSES OF HOMORGANIC AND QUASI-HOMORGANIC V_1##V_2 SEQUENCES

Sylvia Moosmüller

Acoustics Research Department, Austrian Academy of Sciences, Vienna, Austria

ABSTRACT

In the present study, the contribution of the modulation of formant amplitudes for segmenting homorganic or quasi-homorganic V_1##V_2 sequences has been worked out. It can be shown that the modulation of the amplitude of one formant is sufficient to mark the segment boundary. This modulation can either appear as valley at the transition phase or as a sudden rise or fall of the amplitude. The support of a synchronous change in formant frequency is not necessary for segment boundary determination. I.e., even if formant frequencies have already assimilated totally, the segement boundary can still be determined by amplitude modulation. As soon as formant amplitudes start to rise or drop gradually or as soon as they exhibit the same strength over the whole signal, total assimilation has taken place and no segment boundary can be determined any more.

1.INTRODUCTION

Canonically, V_1 ## V_2 sequences should be realized as vowel – glottal stop – vowel, the glottal stop indicating the segment boundary. Speaking both in terms of production and perception, the insertion of the glottal stop has the function of producing maximum contrast. As a function of various linguistc and socio-psychological variables ([4], [5]), reduction and assimilation processes can be observed ([3], [8], [9], [10]). However, the reduction of [V_1ʔV_2] does not follow exactly the same pattern as that which has been described for intervocalic stops ([2], [9]), as the glottal stop in a first step is replaced by creaky voice. Nevertheless, intervocalic sound energy difference [9] remains an important parameter in the process of gradual assimilation. In a previous study [7], the following intermediate steps from canonical realization to total assimilation could be worked out:

1. A glottal stop is substituted by creaky voice.
2. Creaky voice is gradually replaced by modal phonation.
3. The transition phase is marked by a drop in fundamental frequency.
4. Modal phonation, the transition phase is made apparent by a drop of the amplitudes of all formants.
5. The segment boundary is indicated by a sudden rise or drop of the amplitude of one of the formants.
6. The assimilation of formants might take place from stage 1 to stage 5.
7. The assimilation is completed, no indication of a segment boundary can be found.

In this preliminary study, a detailed description of the gradual modulation of formant amplitudes has not yet been given. Consequently it is the aim of the present contribution to give a thorough step by step analysis of stages 4 and 5.

These two stages mark a very progredient phase of the assimilation process, therefore they can be most easily observed when both V_1 and V_2 are completely in unstressed position.

From the first preliminary study it has also been observed that younger people are more prone to apply this assimilation process, consequently, for the present study, spontaneous speech behaviour of younger people has been analysed.

2.METHOD

The spontaneous speech (semi-structured interviews) of five male Viennese Standard speakers (20 – 25 years) has been used [5]. The signal was digitized at 16 kHz, 16 Bit, by means of the Acoustic Workstation S_Tools [1]. The following V_1 ## V_2 sequences have been analyzed:

V_1 ## V_2	canonical transcription	example	translation
er ## a	ɐʔa	immer a	always a
ar ## a	aʔa	war also	was also
er ## e/ei	ɐʔɛ/aɛ	aber es	but it
ar ## e/ei	aʔɛ/aɛ	war eine	was a
e ## a	ɛʔa	Dialekte am	dialects on
a ## e/ei	aʔɛ/aɛ	hamma ein	we have a
e ## e/ei	ɛʔɛ/aɛ	Gefühle erwecken	awake emotions
a ## a	aʔa	da am	there on

Of the sequences the first three formant frequencies and amplitudes were calculated by means of LPC analysis, 22 coefficients and a pre-emphasis 0.9. Measurements were taken at the center of V_1, V_2 and at the (assumed) transition part. In order to evaluate the fine structure of V_1 ## V_2 sequences, frame by frame formant frequency amplitude contours have been produced. The frame width was 64ms and frame shift varied from 2ms to 8ms in order to obtain linear time scale standardization.

3.RESULTS

The starting-point of the present study are cases in which the whole sequence is realized in modal phonation, but the segment boundary is marked by a drop in the amplitude of all three formants at the transition phase. In cases like these, the segment boundary is still distinct, both from the production and perception point of view.

3.1.Segment boundary is marked by amplitude modulation of one formant.

The drop of the formant amplitudes is gradually suspended; so that only one formant amplitude shows a drop at the transition phase, the other two are assimilated. In the following example (figure 1) a sudden drop of the amplitude of F3 can be observed at frame 18, at frame 20 the amplitude rises again. The difference in amplitude amounts to 2dB. The amplitude of the first formant shows the same strength over the whole signal (=total V_1/V_2 assimilation), and the amplitude of the second formant shows a continuous rise, also indicating total V_1/V_2 assimilation. Therefore, the only indicator for a segment boundary is the amplitude valley of F3, but from the perceptual point of view, this amplitude valley is sufficient.

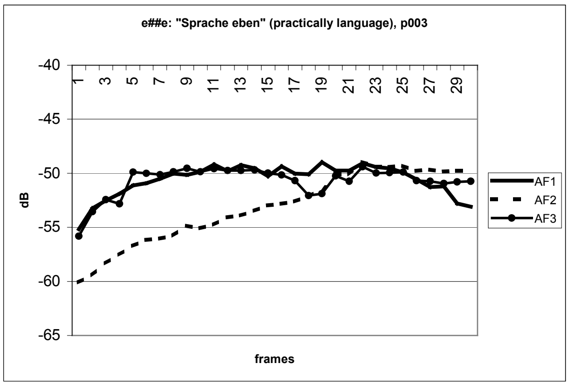


figure 1: frame by frame amplitude contour of F1, F2 and F3

For segmentation, it does not matter at all which formant shows up the amplitude valley. In the next example, the amplitude valley appears during F1 (figure 2):

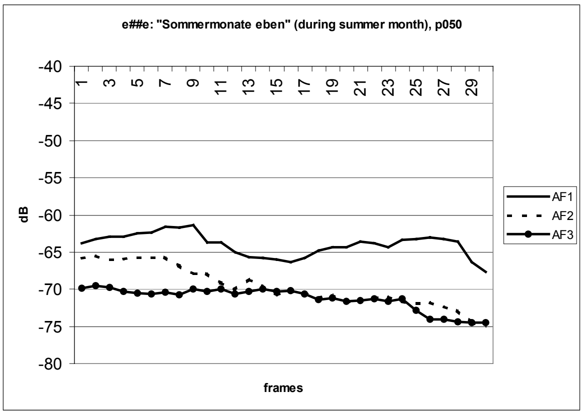
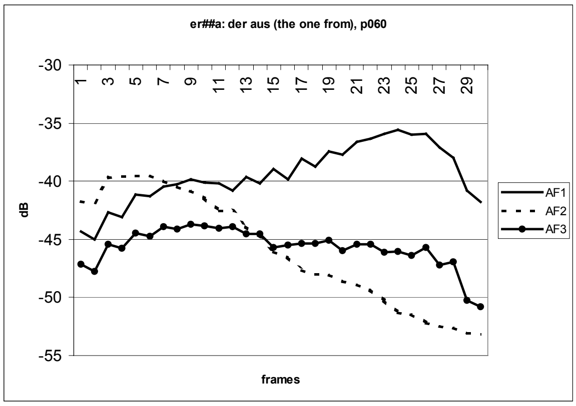


figure 2: frame by frame amplitude contour of F1, F2 and F3

In figure 2, the amplitude valley of F1 has a relatively long duration (frame 9 to frame 18); this would indicate a long and consequently quite blurred transition phase. Therefore, additionally, the drop at frame 9 of F1 is accompanied by a sudden drop in amplitude of F2 at the same point in time. As compared to the previous example, the amplitude of F3 has the same strength over the whole signal (= totally assimilated) in this case.

As assimilation progresses, the transition phase gets gradually lost and the segment boundary is marked by a sudden drop or rise of formant amplitudes. In the following example (figure 3),

figure 3: frame by frame amplitude contour of F1, F2 and F3



the amplitude of F1 suddenly rises at frame 13. The amplitude of F2 continuously drops, therefore, no segment boundary can be determined on the basis of this amplitude. However, the segment boundary indicated at frame 13 is supported by a formant frequency change of F2 and F3 at frame 13 (figure 4):

figure 4: frame by frame formant frequency analysis

However, the support of a formant frequency change is not necessary for the determination of the segment boundary. Even if a total assimilation of formant frequencies has taken place already, a sudden rise or drop of one of the amplitudes suffices completely for determination. In the next example (figure 5), no change in formant frequency over time can be observed, i.e., the signal is totally assimilated with respect to articulatory movement.

However, a sudden drop of the amplitudes of F2 and F3 can be observed at frame 16 (figure 6). This drop indicates the segment boundary. The amplitude of F1 shows a gradual, slow rise and drop and is strongest in the middle of the signal. It can therefore be assumed that a complete V_1/V_2 assimilation with respect to the amplitudes of F1 has taken place.

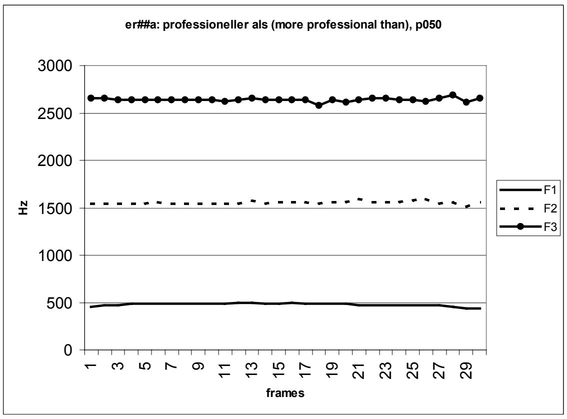


figure5: frame by frame formant frequency change.

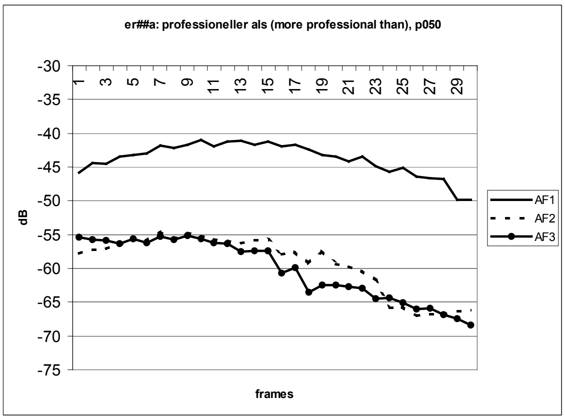


figure 6: frame by frame amplitude contour of F1, F2 and F3

The next example (figure 7) again exposes a total V_1/V_2 assimilation as regards formant frequencies and the only indication of a segment boundary is marked by a sudden drop in

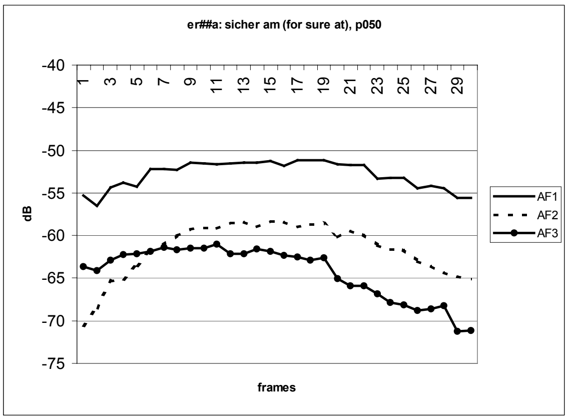


figure 7: frame by frame amplitude contour of F1, F2 and F3

amplitude of F3 at frame 19. Both the amplitudes of F1 and F2 show a gradual, slow rise at the beginning of the signal, followed by a "steady state" and ending with a gradual, slow drop. Consequently, V_1/V_2 assimilation has taken place with respect to the amplitudes of F1 and F2.

3.2.The relevance of formant frequency change

It has become obvious that a change in formant frequencies is not necesssary for the determination of a segment boundary as long as there is a sudden change in one of the formant amplitudes. Therefore, V_1/V_2 formant frequency assimilation might take place earlier than V_1/V_2 amplitude assimilation.

However, the opposite approach might take place as well, although it is less often observed. In the following example (figure 8), the amplitudes of all three formants expose a gradual change, therefore it is not possible to determine a segment boundary on this parameter.

figure 8: frame by frame amplitude contour of F1, F2 and F3.

However, although F2 exposes a rather gradual movement, it starts to change more radically at about frame 20, exhibiting more obviously the values of an [ɛ] (figure 9). Perceptually, the segment boundary can be put here.

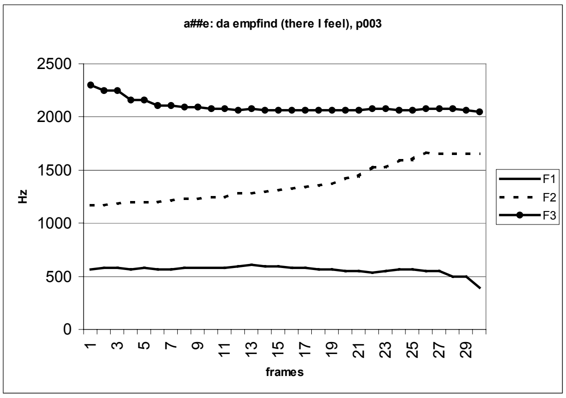


figure 9: frame by frame formant frequency change.

In Austrian German, formant frequency changes usually expose a rather gradual pattern [6], consequently they are not as good a marker for segment boundaries as their amplitudes. This might explain the fact that the V_1/V_2 formant frequency assimilation is allowed to happen at a quite early stage.

An interaction of formant frequency changes and sudden change of one of the formant amplitudes, however, can be observed more often (as an example see figures 3 and 4).

3.3.Total assimilation of V_1 and V_2

V_1##V_2 sequences which no longer expose any sudden change in one of the formant amplitudes must be regarded as having assimilated totally. In cases where only a gradual change can be observed, the segment boundary cannot be determined any more. The following example (figure 10) exposes a gradual drop of the amplitude of F1, the amplitudes of F2 and F3 show up the same strength over the whole sequence.

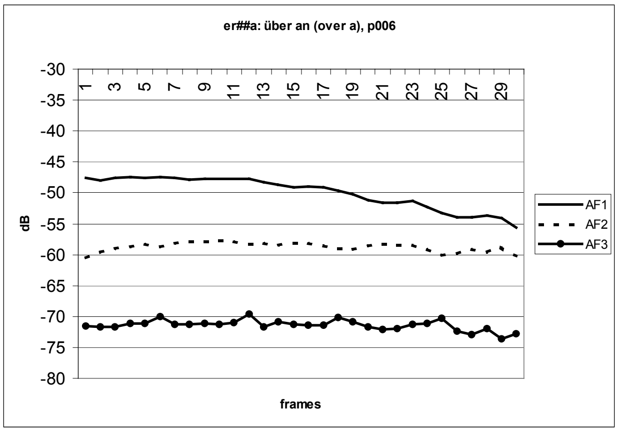


figure 10: frame by frame amplitude contour of F1, F2 and F3

Complete V_1/V_2 formant frequency assimilation has taken place already, perceptually, no segment boundary can be determined any more.

Nevertheless, the process of assimilation is completed when both formant frequencies and formant amplitudes no longer exhibit change over time, as in the last example (figure 11):

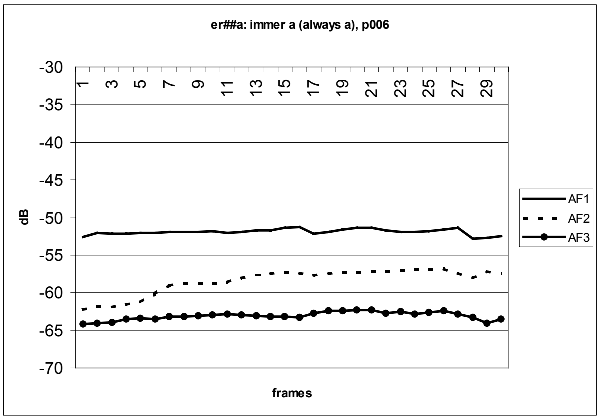


figure 11: frame by frame amplitude contour of F1, F2 and F3

4. CONCLUSION

It follows from these results that the modulation of formant amplitudes is the most important indicator for segment boundaries in V_1##V_2 sequences. In most cases, formant frequency assimilation happens at a quite early stage. Consequently, formant frequency changes give almost no information about the position of a segment boundary.

However, no hierarchy could be detected as to which formant amplitude is assimilated first. Therefore, the amplitudes of all three formants can be used equally for segment boundary indication.

REFERENCES

[1] Deutsch, W. A. and Noll, A. 1994. *S_Tools. Datenerfassung, Speicherung und Digitale Signalverarbeitung für Akustik, Lärm, Sprache und Musik.* Austrian Academy of Sciences, Acoustics Research Department, Vienna.

[2] Duez, D. 1995. On spontaneous French speech: aspects of the reduction and contextual assimilation of voiced stops. *Journal of Phonetics*, 23, 407-427.

[3] Helgason, P. and Kohler, K. J. 1996. Vowel Deletion in the Kiel Corpus of Spontaneous speech. *AIPUK* 30, 115.157.

[4] Madelska, L. and Dressler, W.U. 1996. Postlexical Stress Processes and Their Segmental Consequences Illustrated in Polish and Czech. In: Hurch, B. and Rhodes, R. (eds.), *Natural Phonology: The State of the Arts.* Berlin: de Gruyter, 189-200.

[5] Moosmüller, S. 1991. *Hochsprache und Dialekt in Österreich*. Wien: Böhlau.

[6] Moosmüller, S. 1997. Diphthongs and the process of monophthongization in Austrian German: a first approach. *Proc. EUROSPEECH'97*, 787-790, Rhodos.

[7] Moosmüller, S. 1998. Homorganic and quasi-homorganic V1##V2 sequences in Austrian German. *Sound Patterns of Spontaneous Speech*, 43-47, La Baume-les-Aix.

[8] van Bergem, D. R. 1995. Experimental Evidence for a Comprehensive Theory of Vowel Reduction. *Proc. EUROSPEECH'95*, 1319-1322, Berlin.

[9] van Son, R. J. J. H. and Pols, L. C. W. A Comparison between the Acoustics of Vowel and Consonant Reduction. *IFA Proceedings,* 20, 13-25.

[10] Vollmer, K. 1997. *Koartikulation und glottale Transparenz.* Universität Stuttgart: phonetik AIMS.